Naval Research Laboratory

Washington, DC 20375-5320

NRL/MR/6180--01-8582

Supervisory Control System for Ship Damage Control: Volume 6 — Experimental Measurement of Situation Assessment

Patrick L. Wadlington
Janet A. Sniezek
David C. Wilkins

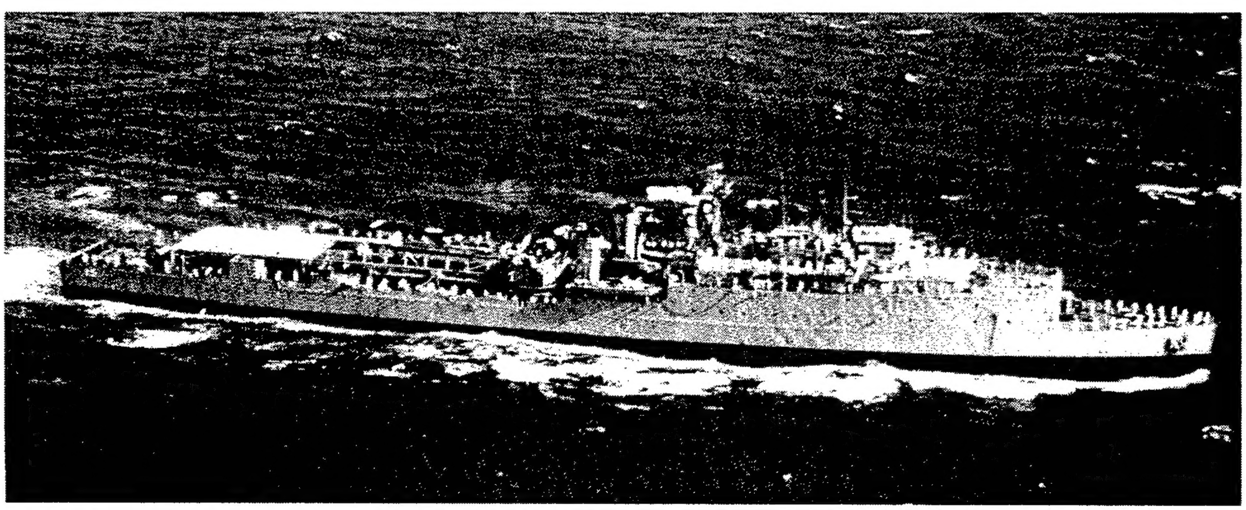
Beckman Institute
University of Illinois, Urbana, Illinois

Patricia A. Tatem
Frederick W. Williams

Navy Technology Center for Safety and Survivability
Chemistry Division

September 28, 2001

20011128 215



REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave Blank*)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 28, 2001	Final FY 1999-FY 2001

4. TITLE AND SUBTITLE

Supervisory Control System for Ship Damage Control: Volume 6 — Experimental Measurement of Situation Assessment

5. FUNDING NUMBERS

PE - 63508N

6. AUTHOR(S)

Patrick L. Wadlington,* Janet A. Sniezek,* David C. Wilkins,* Patricia A. Tatem, and Frederick W. Williams

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Research Laboratory, Code 6180
4555 Overlook Avenue, SW
Washington, DC 20375-5320

8. PERFORMING ORGANIZATION REPORT NUMBER

NRL/MR/6180--01-8582

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5660

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

*Beckman Institute, University of Illinois, Urbana, IL 61801

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

This report develops a methodology for evaluating the degree that a Supervisory Control System for Damage Control (DC-SCS) conveys situation assessment to a human supervisor, who is usually called a Damage Control Assistant (DCA). The first is non-evasive. Based on videotapes of the DC-SCS human-computer interface consoles, a post-session analysis is conducted to determine the extent that the DC-SCS presents the information deemed most relevant to a ship crisis. The second is invasive in that the DCA—whose only source of information on the state of the world is the DC-SCS — is asked questions about state of the ship and the ship crisis.

The construct of situation assessment is described and psychometric issues concerning its accurate measurement are raised. Previous research concerning a model for situation assessment is reviewed. The weaknesses of this past research are determined and recommendations for improvements are made. A new model of situation assessment specific to the damage control domain is presented with a methodology for experimentation and validation.

14. SUBJECT TERMS

Flooding	Damage control	Automation
Fire	Supervisory control	

15. NUMBER OF PAGES

16

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std 239-18
298-102

CONTENTS

SUPERVISORY CONTROL SYSTEM FOR SHIP DAMAGE CONTROL:
VOLUME 6 – EXPERIMENTAL MEASUREMENT OF SITUATION ASSESSMENT

1. Introduction

This report develops a methodology for evaluating the degree to which a Supervisory Control System for Damage Control (DC-SCS) conveys situation assessment to a human supervisor, who is usually called a Damage Control Assistant (DCA). Two methodologies are developed. The first is non-invasive. Based on videotapes of the DC-SCS human-computer interface consoles, a post-session analysis is conducted to determine the extent the DC-SCS presents the information deemed most relevant to a ship crisis. The second is invasive in that the DCA – whose only source of information on the state of the world is the DC-SCS – is asked questions about state of the ship and the ship crisis.

This paper is organized as follows. First, the construct of situation assessment is described and psychometric issues concerning its accurate measurement are raised. Second, previous research concerning a model for situation assessment is reviewed. The weaknesses of this past research are determined and recommendations for improvements are made. Lastly, a new model of situation assessment specific to the damage control domain is presented along with a methodology to be used for experimentation and possible validation.

2. Psychometric Issues and Situation Assessment

This section describes the importance of Situation Assessment and details the major issues that are relevant to its accurate measurement.

2.1 The Construct of Situation Assessment

Situation assessment can be described as possessing the knowledge of what is going on around a person. A more specific scientific definition would be "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" (Endsley & Garland, 2000).

Within the content domain of damage control, the ability to achieve accurate situation assessment is a vital part of effective performance. This has become particularly true with the development of automated control systems. A supervisor (DCA) needs to comprehend not just the state of a crisis (damage control) and the environment (ship and ship environment) but also know the actions being taken by an automated control system (DC-SCS) so that they can be overridden if needed.

A well-developed supervisory control system (DC-SCS) can gain a good picture of what is happening within the situation via a comprehensive sensory detection system. However, just because a DC-SCS presents this information to a supervisor does not necessarily mean that the supervisor has (a) noticed the information or (b) interpreted it correctly and gained the understanding necessary to respond appropriately. Ultimately, a supervisor is responsible for understanding the behavior of a DC-SCS and overriding it when appropriate; therefore, a key

Manuscript approved August 28, 2001.

feature of a DC-SCS is the extent to which it is able to identify salient information and transfer it to the supervisor.

The benefits of objective measurement of the extent to which a DC-SCS system conveys situation assessment are many. In initial design stages, it allows comparison across competing versions of DC-SCS to see how much salient information is actually being relayed to the supervisor by each of them. Furthermore, a valid measure of situation assessment can be a valuable diagnostic tool for identification of shortcomings in any particular DC-SCS. These shortcomings can either be corrected by improving the design of the DC-SCS or be classified as human performance limitations (Hartman & Secrist, 1991). Although human limitations may be alleviated to a certain extent through training, it may be impossible to completely eliminate them (Endsley & Bolstad, 1994).

Accurate measurement of situation assessment, though necessary for damage control success, can be quite elusive. In general, any construct is by definition an unobservable, underlying, casual variable, which can be observed only indirectly through multiple observable behaviors that represent the outcomes of the construct's influence. A person's perception of situation assessment can do more harm than good. Psychometric issues must be taken seriously in the measurement of situation assessment. This requires psychological expertise in the situation assessment domain.

The remainder of this section describes the psychometrics issues most relevant to an accurate situation assessment methodology for the domain of ship damage control.

2.2 Reliability

One of the main psychometric issues is the *reliability* of the situation assessment measure. If a certain situation is measured multiple times, reliable measurements would be nearly, if not completely, equivalent. The difference between the measurements is considered measurement error, which can be detrimental to a supervisor's performance. The validity of a measurement instrument can never exceed the reliability of that instrument.

2.3 Validity

Another major psychometric issue of situation assessment is the *validity* itself. A measurement instrument may be very consistent in repeated measure (i.e., reliable), yet be measuring the construct inaccurately. Many types of validity need to be assessed.

One of the most important types of validity is discriminant validity. This type answers the question, "If two similar situations are measured, can the measurement instrument accurately distinguish between the two?" This question is important because it shows how well a measurement instrument (i.e., SCS) can detect small differences between situations, thereby letting researchers know how discriminating the measurement instrument really is and giving the information needed to improve the measurement instrument or, at least, to recognize its limitations.

Another main type of validity to be assessed is predictive validity. This type answers the question, "Based on the assessment of the situation, can the supervisor make accurate prediction as to how the situation will evolve in the near future?" Its obvious importance is that it allows the measurement instrument to assess how sufficient it is at forecasting future situational events.

2.4 Individual Differences

A third important measurement issue is the difference in performance due to individual differences. How individuals perform based on information given to them by a system varies greatly. This variation can be due to many internal factors of the individual. These factors must be evaluated for three reasons. One reason is to identify the aspects of certain factors (e.g., spatial perceptions) that can be capitalized on in the design and the development of the system. The second reason is for the selection of the best individual for the performance via the system (Baumann, Sniezek, & Buerkle, in press). A third reason is for the assessment of individual weaknesses so they can be trained to overcome them in their performance (Sniezek, Wilkins, & Wadlington, 2000). Individual differences being important, they must be considered in the methodology of evaluating a system's usefulness.

2.5 Supervisory Control System

Within the damage control domain, the measurement of situation assessment is even more complex with the inclusion of a supervisory control system that, independent of the supervisor, takes actions in response to events in the environment. Also, automation serves to increase the complexity of measuring situation assessment. Automation makes the supervisor somewhat blind to aspects of the situation and the system's response. The supervisor may miss the details of the actions of the DC-SCS. This is not bad in itself as long as the macro-actions are portrayed to the supervisor; but, whether this is done or not is a very important question that must be answered. If done, automation may be very beneficial, eliminating micro-tasks with which the system and the supervisor should not be concerned, allowing more of their finite attention to be freed up to deal with the more important issues of damage control. However, if not done, automation creates a lapse in the supervisor's mental model of the situation, which can be extremely detrimental. This being said, the effects of automation on situation assessment must be evaluated with great caution.

A supervisory control system is a whole entity in and of itself that affects both the environment and the supervisor. A supervisory control system receives information from the environment, evaluates the situation using its intelligent reasoning, and generates responses including actions deemed necessary. A supervisory control system also receives environmental information and portrays this information and its own actions as stimuli to the supervisor, who responds accordingly (e.g., possibly overriding). It is impossible to understand a supervisor's situation assessment without taking into account the two aspects of a supervisory control system: the stimuli given by the supervisory control system, and its dynamic interaction with the environment.

One main approach to improving situation assessment is to identify the errors that occur in an individual's mental model of the situation [identifying the omission or commission of relevant (or inappropriate weights there assigned)], to find the source of these errors, and to correct the problem (Durso & Gronlund, 1999; Jones & Endsley, 1996). System assessment allows for the identification of stimuli within the supervisory control system that is giving bad information to the supervisor, thereby causing him/her to evaluate a situation incorrectly. Without putting the supervisory control system itself under its own magnifying glass, many opportunities to eliminate errors and improve a supervisor's situation assessment will be missed.

In evaluating an individual's mental model of a situation, it is important to understand not only what the individual knows, but also the level at which the supervisor understands (Fowlkes, Baker, Cannon-Bawers & Stout, 2000). For instance, the supervisor may know the fire is being fought successfully in a section of a ship. The question is, how successfully? Automation, while beneficial in lessening the mental load on the supervisor, can seriously hinder the depth of understanding of a given situation. Therefore, for maximum supervisor performance, the appropriate fine line needs to be established as to the extent of usage of automation and the way automation information is presented to the supervisor. Automation is a factor that must not be overlooked if any valuable work is to be done with situation assessment in the damage control domain.

A General Model of Situation Assessment

A general breakdown of the components of the situation assessment construct and its interaction with its Individual and Task are present in the following diagram.

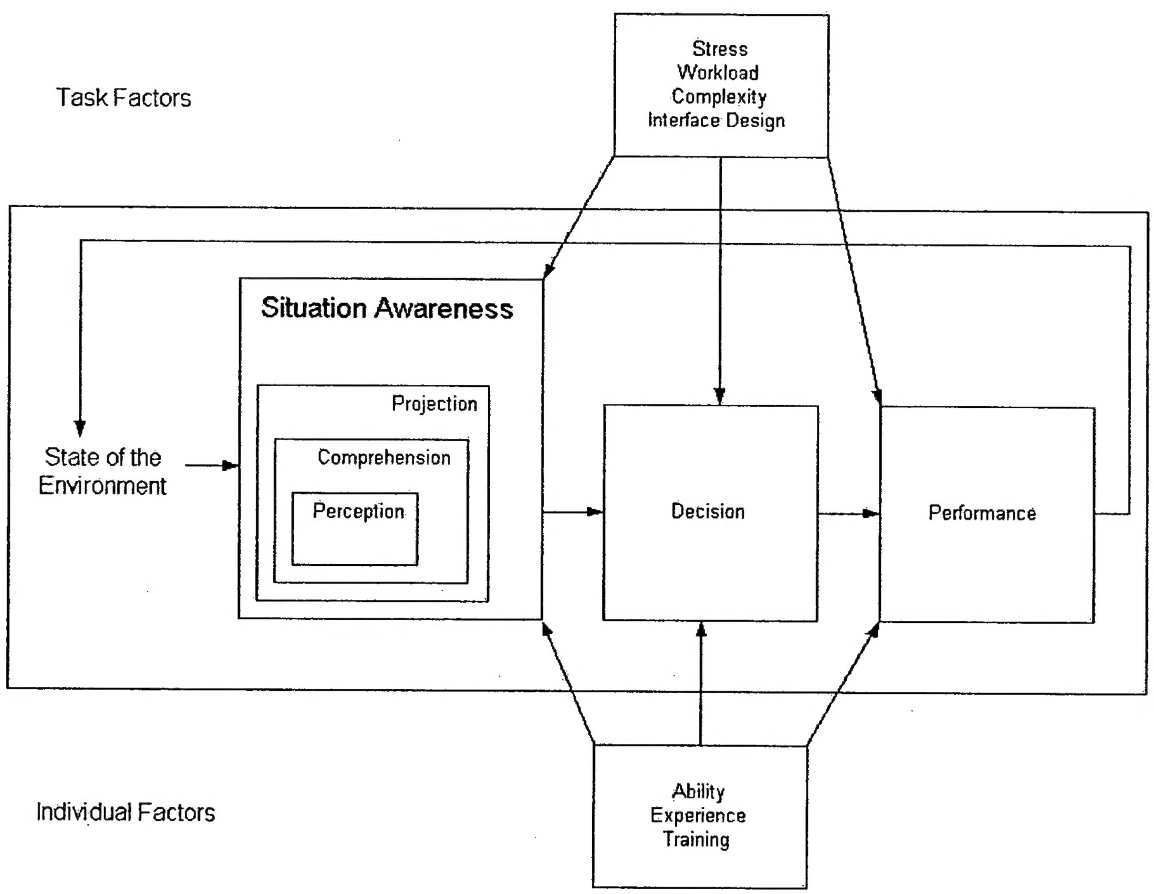


Figure 1. General model of situation assessment

Figure 1 shows a general model of situation assessment. This model is adapted from the situation assessment theory developed by M. R. Endsley (1995). This general model of situation assessment is based on an accumulation of past situation assessment research done primarily in the area of aviation (Durso & Gronlund, 1999; Endsley, 1995; Endsley & Garland, 2000; Hartman & Secrist, 1991; Scerbo & Mouloua, 1999 [Theory]; Endsley & Bolstad, 1994; O'Hare, 1997 [Individual Differences]; Carretta, Perry & Ree, 1996; Fowlkes, Baker, Cannon-Bawers & Stout, 2000; See & Vidulich, 1998 [Utility]; Jones & Endsley, 1996; Vidulich, Stratton, Crabtree & Wilson, 1994; Waag & Houck, 1994 [Measurement]). This basic model contains many of the variables and processes that need to be accounted for when studying situation assessment. It shows three main levels of processing working within the situation assessment construct (Endsley & Garland, 2000).

3.1 Levels of Situation Assessment Processing

Figure 1 shows three different levels of DCA Situation Assessment processing. These will be described in turn.

The lowest level of processing is *perception*. It is simply the identification of individual situational events and individual system actions in response to those events. Aviation studies have shown that approximately three out of four errors in situation assessment are created in this first fundamental level of processes (Jones & Endsley, 1996).

The intermediate level of processing in situation assessment is *comprehension*. Comprehension is the combining, interpreting, storing, and retaining of relevant information that was originally perceived. This middle step is where the individual interprets such attributes as causality among sets of events and actions and deciding which are relevant to his/her present goals.

The highest level of processing in situation assessment is *projection*. It is the highest plane of understanding where the individual grasps the causal connection between all the groups of events and actions and can predict their outcomes. One might say that *projection* is being able to see beyond the trees and appreciate the whole dynamic ecology of the forest. Being able to understand the 'big picture' makes it possible for the individual to forecast future events and dynamics. This forecasting ability makes the individual a most effective decision maker.

3.2 Target of Situation Assessment Processing

In the general situation assessment model, these three levels of processing focus on two main targets, *situational events* and *system responses* (Endsley & Garland, 2000). In the damage control domain, a situational event is any event that causes damage to the naval ship, and a system response is any action issued by the SCS in reaction to a situation event or events. However, the general model does not focus on a third target that contributes greatly to the state of the situation, this being the dynamic interaction between the situational events and system responses. The effects of this interaction can have a great impact on the situation, making it a vital target for the supervisor's assessment of the situation.

3.3 Temporal Reasoning and Situation Assessment Processing

One major dimension of situation assessment not included in this model that is vital to damage control is the important temporal aspect (Endsley & Garland, 2000). According to Endsley and

Gardland, time comes into situation assessment in three distinct ways. First, the perception of time is of obvious importance to situation assessment. An individual needs to be aware of the timing of events and actions. Second, the availability of time to react to a situation is important. An individual needs to know the amount of time the supervisor has before an event occurs or before an action needs to take place. Third, the rate at which valid information becomes outdated as a function of time due to changes in a dynamic environment is needed for the individual to be able to choose what information to trust. This temporal aspect can be a strong point for supervisory control systems due to the near-immediate presentation of environmental information.

One more aspect of time must be considered in the damage control domain, or, in general, in any situation involving advanced technologies. This aspect of time pertains to the dynamic interaction effects of the supervisory control system's responses directed at the environment and how the responses change the environment. These interaction effects can change the volatilities and trends within the situation greatly and must be taken into account

3.4 Limitations of Situation Assessment Processing Model

A great deal of research has been done in the area of situation assessment, mainly in the aviation field. This research can give good theoretical guidance as to understanding situation assessment, but only in a general manner. To use these past findings as a generalization of situation assessment for other domains, especially the domain of damage control, would be quite insufficient. This past research on situation assessment is lacking three main components of the damage control field. One is that the system included in the general model of situation assessment is not a supervisory control system. This factor alone brings up many issues that are not addressed in the general model of situation assessment, such as truly understanding a supervisory control system's probabilistic reasoning, and developing appropriate system trust. Another missing component is the severity of temporal constraints in crises. This element should be studied extensively since a supervisor's situation assessment may leave much to be desired under time pressures. Lastly, even though the automation component of situation assessment has been studied somewhat in the past, it has not been pursued to a satisfactory level. For instance, through automation a supervisor may perceive that a sprinkler system has doused fires in multiple compartments across a section of the ship, but is he actually aware of what this means? Does this mean that the fire has been completely extinguished, or does it only mean that the flames have been subdued, and the temperature has been lowered to a reasonable level for a firefighting team to enter? These two interpretations of the same information call for entirely different behaviors by the supervisor. In this example, the supervisor would need to know the trends of temperature changes to choose the correct actions. Based on these facts, it is safe to say two things are needed: a situation assessment model specifically for the damage control domain, and a unique methodology for evaluation of situation assessment within damage control.

4. Situation Assessment Model for Damage Control

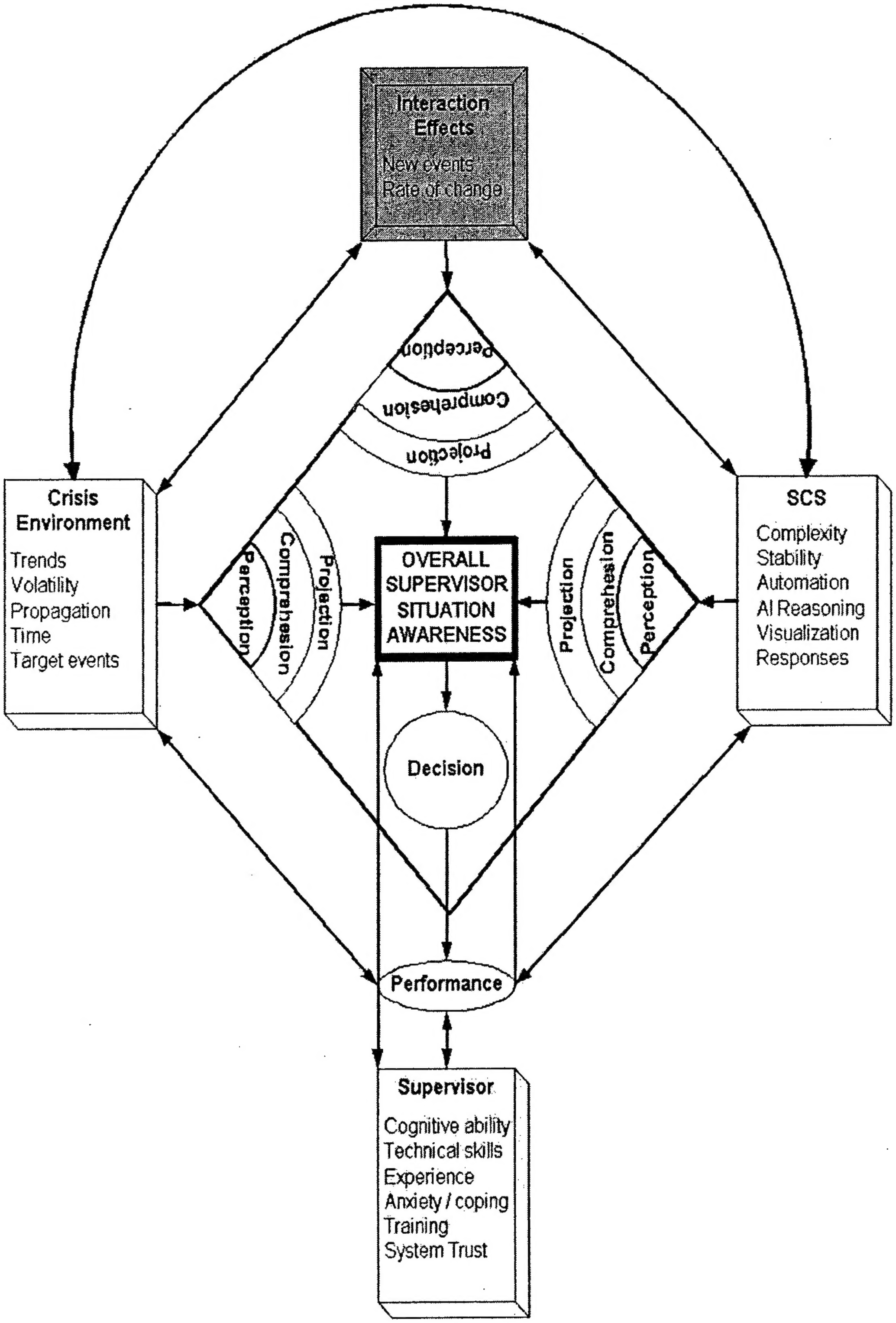


Figure 2. A model of situation assessment for the DC-SCS domain

Figure 2 presents a model of Situation Assessment that we have developed for the domain of ship damage control. This model of situation assessment consists of all vital variables and processes from the general model shown in Figure 1, while eliminating those that do not apply to the damage control domain. In addition, this model contains those factors that are unique and important for domains involving a supervisory control system. This model is made up of four entities: supervisor, crisis environment, supervisory control system (SCS), and interaction effects between the crisis environment and the SCS. Overall supervisor situation assessment consists of the compiled assessments by the supervisor of these four identities across the three levels of processing: perception, comprehension, and projection.

Characteristics of the supervisor to be considered are: his/her cognitive ability, technical skills, experience, anxiety/coping, training, and system trust. These characteristics influence the supervisor's assessment of the other three entities, and also her/his self-perceptions.

Characteristics of the crisis environment to be considered are: trends, volatility, propagation, time, and target events. These characteristics are the key clues as to what is occurring in that hostile setting.

Characteristics of the supervisor control system to be considered are: complexity, stability, automation, artificial intelligence (AI), reasoning, visualization, and responses. These characteristics allow the supervisor to judge the performance of the SCS and establish his/her level of trust in the system which determines her/his intentions as to overriding the SCS with a manual command.

Characteristics to be considered in interaction effects between the crisis environment and the SCS are: new events and rate of change in the associated entities. Taking these interaction effects into account makes the projection process much more feasible.

Having taken these four entities into assessment, the supervisor gains his/her overall supervisor situation assessment. Then she/he makes decisions based on this assessment and takes action (performance) accordingly. These actions in turn change the state of all four entities and the overall supervisor situation awareness.

4.1 Situation Assessment Checklist

Type of SCS:__________________ Date of appraisal:_______ Appraiser:____________

Damage Awareness Type	Yes/No, To what extent	Time of Awareness
Identification of individual target events (Perception)		
Fire		
compartment location		
location within compartment		
type of fire		
severity of fire		
status of fire		
rate of spread		
how long been ignited		
Flooding		
compartment location		
severity of flooding		
status of flooding		
rate of spread		
how long been flooding		
Identification of individual status information (Perception)		
Firemain status		
Door closure positions		
Ventilation suppression		
Fire suppression		
Identification of resources (Perception)		
Men		
repair parties		
how many		
location of each		
flex teams		
how many		
location of each		
Identification of system responses to target events (Perception)		
Water mist		
Fire boundaries status		
Teams sent		
Assessment of causality in a set of system responses (Comprehension)		
Assessment of grouped target events (Comprehension)		
Global understanding of overall situation (Projection)		
Accuracy of trust (supervisor override calibration)		
Level of confidence (certainty calibration)		

Figure 3. Situation assessment checklist

The situation assessment checklist (Figure 3) is a major part of the observational research. It is to be filled out by a University of Illinois (UIUC) experimenter or another designed impartial experimenter proficiently trained to carry out the standardized methodology. The checklist is an exhaustive account of the mental model that can be gained from the SCS via direct observation of the SCS and via the supervisor's performance.

4.2 Evaluation of DC-SCS Using Situation Assessment Model

To address this model's implications, we have developed a unique and specific approach to the evaluation of situation assessment within the specific domain of damage control. This distinct methodology for situation assessment evaluation was developed in parallel with the DC-SCS itself, allowing the UIUC researchers to have the distinct advantage of creating DC-SCS with situation assessment in mind. This hallmark methodological approach toward situation assessment will continue to benefit DC-SCS, allowing for continued improvements in the potential situation assessment gained by the supervisor across a variety of crisis situations.

In addition, the UIUC approach allows *any* researcher to assess the limits of situation assessment in and across all SCS systems and supervisors. In any specific SCS, this methodology can be used to identify faults, correctable or not, that delay or diminish the potential for situation assessment by the supervisor. Across versions of SCS, this methodology allows researchers to find limitations and/or difficulty levels within the attributes of scenarios, such as identifying the salient information that is most difficult for a supervisor to grasp and the non-salient information that is most disruptive across the three levels of processes in the gaining of situation assessment. For the supervisors themselves, this methodology can expose the limitations within their personal attributes for gaining certain competencies out of the potential situation assessment to which they are exposed. This latter use of the methodology can be helpful in identifying the focus of needs in personnel training and in regards to the criterion used for personnel selection (Endsley & Bolstad, 1994; O'Hare, 1997).

4.3 Facets of Situation Assessment Evaluation

The UIUC approach to situation assessment evaluation can be broken down into three main facets. It is of great importance to measure situation assessment with multiple method facets; otherwise, measurement error may be mistaken for a significant difference (Vidulich, Stratton, Crabtree & Wilson, 1994). The first facet of evaluation is the systematic observation and measurement of supervisor behaviors. This facet will be measured in accordance with system performance. Supervisors and domain experts will be the personnel evaluated. They will be evaluated via the simulator of DC-SCS. The second facet of situation assessment evaluation is actual testing done on the Ex-USS Shadwell, which will be preceded by pilot testing on shore to ensure the maximizing of research efforts while aboard the ship. The last facet of evaluation is objective and descriptive reports on supervisors' performance. This will be conducted with self-report and evaluative surveys done by supervisors and domain experts, respectively.

4.4 Study 1 -- Observational Research

The UIUC approach to situation assessment evaluation focuses on two main areas of research. The first focus is on observational research. This research focus is very non–invasive. Without interfering in the Navy's daily business, UIUC researchers are able to narrow down which concepts of situation assessment are most important and direct the Navy's energies accordingly. This allows the identification of important dimensions of the situation assessment construct without interfering in the natural development of the phenomenon; however, it does not allow intensive study of any one of these dimensions. This is where the second main focus comes into play with dimension-specific research. Dimension-specific research is invasive in that it does interfere in the ongoing performance of the supervisor. It is control intensive and allows

research to uncover critical issues such as causality through the manipulation of relevant variables.

The methodology developed for observational research involves a comparison of the SCS's presentation, objective measures of the supevisor, and the test characteristics of each specific scenario. The SCS's presentation will be recorded via video recorder. The video will record the comprehensive information needed concerning each specific scenario and time frame across the visualization, interface, and sound. To make the comparison feasible, this presentation data will be scored to determine the properties of the salient information elements within the SCS, such as the number of salient data pieces presented out of total salient data pieces that exist and the number of salient data pieces presented out of all data pieces presented. Also, ease of identification will be determined from the accessibility (alarm, interface, menu option), the type of stimuli (audio, visual), and the amount of interference (clutter).

The objective measures of the supervisor's performance will also be analyzed from video output. The supervisor objective measures are correct actions, errors made, and response times. They can be readily used in a comparison without any manipulation procedure. The test characteristics of each specific scenario will be taken from the comprehensive testing document produced by NRL. The test characteristics to be used in the comparison are the accounts of the events in each scenario, the appropriate assessments of situation (not via the SCS), and the appropriate responses to the events of the crisis situation.

After these three sources of information have been collected, all the information gained from the SCS's presentation and the supervisor will be compiled into an all-possible information checklist for situation assessment. Finally, this comprehensive checklist of the supervisor and SCS's presentation will be compared to all the data collected on the test characteristics for each scenario.

The results of the comparison in this observational research will be very useful. First, they will allow researchers to find all the discrepancies between SCS/supervisor mental model of the situation assessment and the appropriate mental model intended by test characteristics. The discrepancies can be identified across all levels of analysis including the presentation of the situation by SCS, the identification of individual target events by supervisor, the assessments of groups of target events by supervisor (via individual supervisor responses), and the understanding of causality of global situation (via overall supervisor performance scores).

A second advantage of having these results is that they make it possible for researchers to eliminate any non-salient information (Kass, Herschler & Companion, 1991). The elimination of this 'clutter' is a major request from damage control naval personnel. One benefit of eliminating this information is that it allows for more of the supervisor's attention to be free to take account of the salient information presented. In addition, from a pragmatic standpoint, this elimination of clutter frees up space on the interface for the addition of previously unaccounted-for salient information. Lastly, these results allow for the identification of additional salient information, which can be added so as to continue the improvement and refinement of an SCS that maximizes the potential situation assessment of a supervisor.

4.2 Study 2 -- Dimension-Specific Research

The methodology developed for dimension-specific research consists of four procedures. Before the first procedure, general instructions and system training are given to the supervisors. It

includes an interface introduction (task description, interface familiarity) and a measurement of baseline reaction times to mock events and SCS actions.

Testing techniques make up the first procedure for the dimension-specific research. These testing techniques are the measurement of actual reaction times (individual events, individual actions), identification tests via task interruption, and prediction tests via task freezing at one, ten, and thirty minute intervals. The measurement of reaction times gives researchers an accurate interpretation of how long it takes for a supervisor to assess the situation and response accordingly. Identification tasks allow researchers to assess the degree of accuracy in the supervisor's situational assessment and her/his depth of understanding at any point throughout the situation. The prediction tests convey the accuracy to which a supervisor can forecast future events over different durations of time based on his/her situation assessment.

The second procedure consists of the events to be identified. They are the location of the crisis (i.e., fire and primary damage areas) and the status of primary functions (fire main, door closure positions, ventilation, fire suppression, SCS actions).

System actions to be identified as a function of time are the third procedure. They are divided into actions directed toward controlling areas inside the primary damage areas (PDA) and those controlling areas outside the PDA. The actions directed within the PDA are those for controlling the fire effects (setting vertical boundaries, setting horizontal boundaries, recovering twenty feet visibility within compartments and passageways, controlling extraneous fire) and those for controlling the fire main (controlling flooding, cooling boundaries, protecting magazines, cooling water, dewatering, washing down the exterior). The actions directed outside the PDA are those for controlling the fire main (controlling fire, extinguishing fire) and those for the controlling the fire (controlling flooding, cooling boundaries, protecting magazines, cooling water, dewatering, washing down the exterior).

The fourth and final procedure for the dimension-specific research consists of two evaluations. The first evaluation is to be completed by supervisors. It will help to identify supervisors' level of trust of the SCS, which is of great importance when it comes to gaining the acceptance of a SCS and calibrating the supervisory override function. The second evaluation is to be filled out by domain experts. This evaluation will similarly assess the level of trust of the SCS by the experts but it will also assess the experts' level of trust in the supervisors. This second assessment is another way of evaluating whether the salient information is being perceived by the supervisor via the SCS.

5. Summary and Conclusions

This report developed a methodology for evaluating the degree that a Supervisory Control System for Damage Control (DC-SCS) conveys situation assessment to a human supervisor, who is usually called a Damage Control Assistant (DCA).

Through this methodology, the level of situation assessment that a DC-SCS provides to a supervisor can be maximized, thus maximizing the effectiveness of the joint reasoning of the DC-SCS and DCA cooperating as a team.

6. References

Baumann, M. R., Sniezek, J.A., & Buerkle, C. A. (in press). *Self-evaluation, stress, and performance: a model of decision making under acute stress.* Book Chapter in Klein & Salas (Eds.) NDM 4.

Carretta, T. R., Perry Jr., D. C., & Ree, M. J. (1996). *The International Journal of Aviation Psychology, 6*(1), 21-41.

Durso, F. T. & Gronlund, S. D. (1999). *Handbook of Applied Cognition.* West Sussex, England: John Wiley & Sons Ltd.

Endsley, M. R. (1995). Toward a theory of situation assessment in dynamic systems. *Human Factors, 37*(1), 32-64.

Endsley, M. R., & Bolstad, C. A. (1994). Individual differences in pilot situation assessment. *The International Journal of Aviation Psychology, 4*(3), 241-264.

Endsley, M. R., & Garland, D. J. (2000). *Situation Assessment Analysis and Measurement.* Mahwah, NewJersey: Lawerance Erlbaum Associates, Inc.

Fowlkes, J. E., Baker, D. P., Cannon-Bawers, J. A., & Stout, R. J. (2000). The utility of event-based knowledge elicitation. *Human Factors, 42*(1), 24-35.

Hartman, B. O., & Secrist, G. E. (1991). Situational assessment is more than exceptional vision. *Aviation, Space, and Environmental Medicine, 62,* 1084-1089.

Jones, D. G., & Endsley, M. R. (1996). Sources of situation assessment errors in aviation. *Aviation, Space, and Environmental Medicine, 67,* 507-512.

Kass, S. J., Herschler, D. A., & Companion, M. A. (1991). Training situational assessment through pattern recognition in a battlefield environment. *Military Psychology, 3*(2), 105-112.

O'Hare, D. (1997). Cognitive ability determinants of elite pilot performance. *Human Factors, 39*(4), 540-552.

Scerbo, M. W., & Mouloua, M. (1999). *Automation Technology and Human Performance.* Mahwah, NewJersey: Lawerance Erlbaum Associates, Inc.

See, J. E., & Vidulich, M. A. (1998). Computer modeling of operator mental workload and situational assessment in simulated air-to-ground combat: an assessment of predictive validity. *The International Journal of Aviation Psychology, 8*(4), 351-375.

Sniezek, J. A., Wilkins, D. C., & Wadlington, P. L. (2000). *Advanced Training for Crisis Decision Making: Simulation, Critiquing, and Immersive Interfaces.* Presentation given at HCISS 34th Annual Conference.

Vidulich, M. A., Stratton, M., Crabtree, M., & Wilson, G. (1994). Performance-based and physiological measures of situational assessment. *Aviation, Space, and Environmental Medicine, 65*(5), A7-A12.

Waag, W. L., & Houck, M. R. (1994). Tools for assessing situational assessment in an operational fighter environment. *Aviation, Space, and Environmental Medicine, 65*(5), A13-A19.